Attorney Docket No.13757US03

SYSTEM AND METHOD FOR VIDEO AND AUDIO ENCODING ON A SINGLE CHIP

RELATED APPLICATIONS

[01] This application is a continuation-in-part of United States Patent Application Serial No. 10/170,019 filed June 11, 2002, which is a continuation-in-part of United States Patent Application Serial No. 09/543,904 filed April 6, 2000, which claims the benefit of Israel Application Serial No. 129345 filed April 6, 1999.

[02] This application also makes reference to, claims priority to and claims the benefit of United States Provisional Patent Application Serial No. 60/296,766 filed on June 11, 2001 and United States Provisional Patent Application Serial No. 60/296,768 filed on June 11, 2001.

[03] All of the above-listed patent applications are incorporated herein by reference in their entirety.

BACKGROUND OF THE INVENTION

[04] Methods for encoding an audio-visual signal are known in the art. According to the methods, a video signal is digitized, analyzed and encoded in a compressed manner. The methods are implemented in computer systems, either in software, hardware or combined software-hardware forms.

[05] Most hardware encoding systems consist of a set of semiconductor circuits arranged on a large circuit board. State of the art encoding systems include a single semiconductor circuit. Such a circuit is typically based on a high-power processor.

[06] Reference is now made to Fig. 1, which is a block diagram illustration of a prior art video encoding circuit 10.

[07] Encoding circuit 10 includes a video input processor 12, a motion estimation processor 14, a digital signal processor 16 and a bitstream processor 18. Processors 12-18, respectively, are generally connected in series.

[08] Video input processor 12 captures and processes a video signal, and transfers it to motion estimation processor 14. Motion estimation processor 14 analyzes the motion of the video signal, and transfers the video signal and its associated motion analysis to digital signal processor 16. According to the data contained within the associated motion analysis, digital signal processor 16 processes and compresses the video signal, and transfers the compressed data to bitstream processor 18. Bitstream processor 18 formats the compressed data and creates therefrom an encoded video bitstream, which is transferred out of encoding circuit 10.

[09] It will be appreciated by those skilled in the art that such an encoding circuit has several disadvantages. For example, one disadvantage of encoding circuit 10 is that bitstream processor 18 transfers the encoded video bitstream, data word by data word, directly to an element external to encoding circuit 10. Accordingly, each time such data word is ready, the encoded video data word is individually transferred to the external element. Transfer of the encoded video in such a fashion requires the use of a special external element to store the

data before it will be transferred via a computer bus, for example, to a storage element or computer memory. Additionally, circuit 10 requires a dedicated storage/bus which is allocated on a full time basis, hence, magnifying these disturbances.

[10] Another disadvantage is that encoding circuit 10 is able to perform the encoding of video signals, only. Usually, moving picture compression applications include multiframe videos and their associated audio paths. While the encoding circuit 10 performs video compression and encoding, the multiplexing of compressed video, audio and user data streams are performed separately. Such an approach increases the data traffic in the compression system and requires increased storage and processing bandwidth requirements, thereby greatly increasing the overall compression system complexity and cost.

[11] Reference is now made to Fig. 2, which is a block diagram of a prior art video input processor 30, as may be typically included in encoding circuit 10. Video input processor 30 includes a video capture unit 32, a video preprocessor 34 and a video storage 36. The elements are generally connected in series.

4

[12]   Video capture unit 32 captures an input video signal and transfers

it to video preprocessor 34. Video preprocessor 34 processes the video

signal, including noise reduction, image enhancement, etc., and transfers

the processed signal to the video storage 36. Video storage 36 buffers

the video signal and transfers it to a memory unit (not shown) external to

video input processor 30.

[13]   It will be appreciated by those skilled in the art that such a video

input   processor   has   several   disadvantages.   For   example,   one

disadvantage of processor 30 is that it does not perform image resolution

scaling. Accordingly, only original resolution pictures can be processed

and encoded.

[14]   Another disadvantage is that processor 30 does not perform

statistical analysis of the video signal, since in order to perform

comprehensive statistical analysis a video feedback from the storage is

necessary, thus allowing interframe (picture to picture) analysis, and

processor 30 is operable in "feed forward" manner, only.

[15]   Reference is now made to Fig. 3 which is a block diagram

illustration of a prior art video encoding circuit 50, similar to encoding

5

circuit 10, however, connected to a plurality of external memory units. As an example, Fig. 3 depicts circuit 50 connected to a pre-encoding memory unit 60, a reference memory unit 62 and a post-encoding memory unit 64, respectively. Reference is made in parallel to Fig. 4, a chart depicting the flow of data within circuit 50.

[16] Encoding circuit 50 includes a video input processor 52, a motion estimation processor 54, a digital signal processor 56 and a bitstream processor 58. Processors 54 to 58, respectively, are generally connected in series.

[17] In the present example, video encoding circuit 50 operates under MPEG video/audio compression standards. Hence, for purposes of clarity, reference to a current frame refers to a frame to be encoded. Reference to a reference frame refers to a frame that has already been encoded and reconstructed, preferably by digital signal processor 56, and transferred to and stored in reference memory unit 62. Reference frames are compared to current frames during the motion estimation task, which is generally performed by motion estimation processor 54.

[18]   Video input processor 52 captures a video signal, which contains a current frame, or a plurality of current frames, and processes and transfers them to external pre-encoding memory unit 60.   External pre-encoding memory unit 60 implements an input frame buffer (not shown) which accumulates and re-orders the frames according to the standard required for the MPEG compression scheme.

[19]   External pre-encoding memory unit 60 transfers the current frames to motion estimation processor 54.   External reference memory unit 62 transfers the reference frames also to motion estimation processor 54. Motion estimation processor 54, reads and compares both sets of frames, analyzes the motion of the video signal, and transfers the motion analysis to digital signal processor 56.

[20]   Digital signal processor 56 receives the current frames from the external pre-encoding memory 60, and according to the motion analysis received from motion estimation processor 54, processes and compresses the video signal.   Digital signal processor 56 then transfers the compressed data to the bitstream processor 58.   Digital signal processor 56 further reconstructs the reference frame and stores it in

7

reference memory 62. Bitstream processor 58 encodes the compressed data and transfers an encoded video bitstream to external post-encoding memory unit 64.

[21] It will be appreciated by those skilled in the art that such an encoding circuit has several disadvantages. For example, one disadvantage of encoding circuit 50 is that a plurality of separate memory units are needed to support its operations, thereby greatly increasing the cost and complexity of any encoding system based on device 50. The three memory units described above (pre-encoding, reference and post-encoding) may be part of one external memory (i.e, placed in the same memory), and then the cost and complexity of the encoding system is not increased greatly. However, the use of a single memory with three parts does not permit simultaneous access to each part, thereby slowing down the decoding process.

[22] Another disadvantage is that encoding circuit 50 has a plurality of separate memory interfaces. This increases the data traffic volume and the number of external connections of encoding circuit 50, thereby greatly increasing the cost and the complexity of encoding circuit 50. As

8

mentioned above, the three memory units described above (pre-encoding, reference and post-encoding) may be part of one external memory (i.e., placed in the same memory), and then the same memory interface may be used, in which case the cost and complexity of the encoding system is not increased greatly. Again, however, the use of a single memory with three parts does not permit simultaneous access to each part, thereby slowing down the decoding process. Another disadvantage is that encoder circuit 50 does not implement video and audio multiplexing, which is typically required in compression schemes.

[23] Further limitations and disadvantages of conventional and traditional approaches will become apparent to one of skill in the art, through comparison of such systems with embodiments of the present invention as set forth in the remainder of the present application with reference to the drawings.

BRIEF SUMMARY OF THE INVENTION

[24] Certain embodiments of the present invention provide an apparatus for performing video and audio encoding. In particular, certain embodiments provide for performing video and audio encoding on a single chip.

[25] Apparatus of the present invention provides for performing real time video/audio encoding on a single chip. Within the single chip, a video encoder generates encoded video data from uncompressed video data and an audio encoder generates encoded audio data from uncompressed audio data. A mux processor within the single chip generates an output stream of encoded data from the encoded video data and the encoded audio data.

[26] These and other advantages and novel features of the present invention, as well as details of an illustrated embodiment thereof, will be more fully understood from the following description and drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[27]   Fig. 1 is a block diagram of a prior art video encoding circuit.

[28]   Fig. 2 is a block diagram of a prior art video input processor.

[29]   Fig. 3 is a block diagram of a prior art video encoding circuit linked to a plurality of external memory units.

[30]   Fig. 4 is a flow chart of the data flow within the prior art circuit illustrated in Fig. 3.

[31]   Fig. 5 is a block diagram of a video and audio encoding video/audio/data multiplexing device constructed and operative on a single chip in accordance with an embodiment of the present invention.

[32]   Fig. 6 is a detailed block diagram of a PCI interface of the device of Fig. 5 in accordance with an embodiment of the present invention.

[33]   Fig. 7 illustrates a block diagram of a I2C/GPIO interface of the device of Fig. 5 in accordance with an embodiment of the present invention.

[34]   Fig. 8 is a block diagram and timing diagram illustrating the signals and timing output by a DVB formatter of the device in Fig. 5 in accordance with an embodiment of the present invention.

[35]   Fig. 9 illustrates how a VBI extractor of the device in Fig. 5 may extract user data from specified lines of a video signal in accordance with an embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

[36]   An embodiment of the present invention provides a video/audio encoder on a single chip to generate compressed video and audio multiplexed into different types of streams (VES, AES, program, transport and other user defined). One embodiment of the encoder of the present invention supports MPEG-1 and MPEG-2 standards and AC-3 standards, for example. Applications for the encoder of the present invention may include personal video recorders, DVD recorders, set top box recorders, PC TV tuners, digital camcorders, video streaming, video conferencing, and game consoles.

[37]   Reference is now made to Fig. 5, a block diagram of video encoding video/audio/data multiplexing device 100, constructed and operative in accordance with an embodiment of the present invention.

[38]   An embodiment of the present invention overcomes the disadvantage of the prior art by providing a novel approach to video/audio compression and encoding, and, as per this approach, a novel encoding device structure which comprises a plurality of processors with a defined, optimized work division scheme.

[39]   Typically, a sequence of compression commands are instructions or a sequence of instructions, such as, removal of temporal redundancy, removal of spatial redundancy, and entropy redundancy of data, and the

like. Device 100 operates according to an optimized compression labor division, thus segmenting the compression tasks between the different processors and reducing, in comparison to prior art, the compression time.

[40] According to an embodiment of the present invention, device 100 is a parallel digital processor implemented on a single chip and designed for the purposes of real-time video/audio compression and multiplexing, MPEG-1 and MPEG-2 encoding. For purposes of clarity herein, multiplexing refers to the creating of synchronized streams of a plurality of unsynchronized audio and video streams. Device 100 may be incorporated in digital camcorders, recordable digital video disk (DVD), game machines, desktop multimedia, video broadcast equipment, video authoring systems, video streaming and video conferencing equipment, security and surveillance systems, and the like.

[41] According to an embodiment of the present invention, device 100 efficiently performs video compression tasks such as removing temporal redundancy (i.e., motion between frames), spatial redundancy (i.e. motion within frame), and entropy redundancy of data. Device 100 has a plurality of processors, each processor designed to perform a segment of the compression task, hence, achieving optimal performance of each such task.

[42] The number of processors, the architecture of each processor, and the task list per processor, achieves the optimal tradeoff between device implementation cost and efficiency.

[43] In an embodiment of the present invention, device 100 incorporates both video encoding and audio encoding on a single chip. Device 100 includes a video input buffer (VIB) 102, a global controller 104, motion estimation processors P4 105 and MEF 106, a digital signal processor (DSP) 108, a memory controller 110, a bitstream processor (BSM) 112, an audio encoder (AUD) 113, a multiplexing processor (MUX) 114, a PCI interface 115, and a I2C/GPIO interface 116.

[44] Together, the VIB 102, MEF 106, P4 105, DSP 108, and BSM 112 constitute a video encoder in an embodiment of the present invention.

[45] Device 100 may be connectable to an external video interface, an external audio interface, an external memory unit, and an external host interface. Typically, for example, the video interface supplies a digital video signal in CCIR 656 format and the audio interface supplies a digital audio signal in I2S/AC97 formats.

[46] The host interface typically connects to an external host (not shown) and acts as a user interface between device 100 and the user. The host interface accepts microcodes, commands, data parameters and the like received from a user or a supervising system. The host interface also may be used to transfer information from device 100 to the user. The host interface provides access to the compressed data and may be

14

used to transfer uncompressed digitized video and/or audio and/or user data into device 100.

[47] The PCI interface 115 connects the single chip device 100 to a PCI bus for use in PC applications. Using the PCI interface 115, the device 100 may directly communicate with the PCI bus without the aid of an intermediate interface (chip) external to the device 100. In an embodiment of the present invention, the heart of the PCI interface 115 includes a powerful programmable DMA engine that may transfer encoded data from the device 100 to host memory without a host processor intervening. Fig. 6 is block diagram of an embodiment of the PCI interface 115 including a PCI core 120, a PCI application 121, and a host interface controller 122. The PCI core 120 provides the interface between the PCI bus and the PCI application 121. The PCI application interfaces the PCI core 120 to the host interface controller 122 and is responsible to the Master/Slave protocols and to configure PCI memory space. The PCI application 121 also includes the programmable DMA engine for transferring compressed data to Host memory. All microcodes and user defined parameters are uploaded to the single chip device 100 through the host interface controller 122 (off-line, prior to operation).

[48] In an embodiment of the present invention, the PCI interface 115 may also support a file mode where an uncompressed file may be brought into the single chip device 100 and encoded. For example, video or/and audio files stored on a PC may be converted to MPEG-2 using this

method. The PCI interface 115 allows the uncompressed file to be transferred quickly to the device 100.

[49] In an embodiment of the present invention, device 100 is operable either in a programming mode or an operational mode, and is capable of operating in both modes simultaneously.

[50] In the programming mode, an external host transfers, via the host interface, commands and data parameters to global controller 104. Global controller 104 transfers the commands and data parameters to video input buffer 102, motion estimation processors 105 and 106, digital signal processor 108, memory controller 110, bitstream processor 112, I2C/GPIO interface 116, and multiplexing processor 114.

[51] In the operational mode, video input buffer 102 is responsible for acquiring an uncompressed CCIR-656 video signal from an external video source (not shown) and storing it via the memory controller 110. In an alternative embodiment, VIB 102 captures an uncompressed video signal, via the PCI interface 115. VIB 102 is responsible for acquiring an uncompressed CCIR-656 video and storing it via the memory controller 110 in an external memory unit in a raster-scan manner.

[52] In an embodiment of the present invention, the memory controller 110 is a SDRAM controller and the external memory unit is an SDRAM memory unit. The SDRAM controller is responsible for communication between the single chip and the external SDRAM memory unit, which is used as a frame buffer and an output buffer for compressed data. The

16

SDRAM controller operations are controlled and scheduled by special instructions issued by the global controller 104.

[53] Video input buffer 102 performs statistical analysis of the video signal, thereby detecting developments in the video contents, such as scene change, sudden motion and the like. Video input buffer 102 also performs horizontal resolution down-scaling, thereby allowing or enabling compression not only of the original resolution frames, but also reduced resolution frames (such half D1 etc.). Additionally, video input buffer 102 also pre-processes the video signal, such as spatial filtering, and the like.

[54] Video input buffer 102 accumulates a portion of the scaled and processed video data and transfers the data in bursts to an external memory unit, via memory controller 110. Memory controller 110 stores the video data in the external memory unit.

[55] In an embodiment of the present invention, device 100 operates under MPEG-1 and MPEG-2 video/audio compression standards. Hence, a data block represents a macroblock, which is a sixteen by sixteen matrix of luminance pixels and two, four or eight, by eight matrices of chrominance pixels as defined by MPEG standards. For purposes of clarity herein, reference to a reference frame refers to a frame that has already been encoded, reconstructed and stored in an external memory unit, and which is compared to the current frame during the motion estimation performed by motion estimation processors 105 and 106.

17

[56]   Motion estimation processor 105 (P4) is a level 1 motion estimation engine that is responsible for downscaling current and original reference pictures and for motion vector search. Motion estimation processor 105 finds motion vectors with a 2-pel accuracy by applying a fully exhaustive search in the range of +/- 96 pels horizontally and +/- 64 pels vertically.

[57]   Motion estimation processor 106 (MEF) is a level 2 motion estimation engine that is responsible for finding final (half pel) motion vectors.   Additionally,   the   MEF   performs   horizontal   and   vertical interpolation of a chrominance signal.   The MEF employs a fully exhaustive search in the range of +/-1 pel (incorrect statement) horizontally and vertically. After the full-pel motion vector is found, the MEF performs half-pel motion search in eight possible positions surrounding the optimal full-pel vector.

[58] The dual memory controller 110 retrieves a current frame macroblock, and certain parts of the reference frames (referred hereto as search area) from the external memory unit and loads them into motion estimation processors 105 and 106. The motion estimation processors compare the current frame macroblock with the respective reference search area in accordance with a sequence of compression commands, thereby producing an estimation of the motion of the current frame macroblock. The estimation is used to remove temporal redundancy from the video signal.

[59] Motion estimation processors 105 and 106 transfer the resulting motion estimation to global controller 104. Motion estimation processors 105 and 106 also transfer the current frame macroblock and the corresponding reference frames macroblocks to digital signal processor 108.

[60] Digital signal processor 108 performs a series of macroblock processing operations intended to remove the spatial redundancy of the video signal, such as discrete cosine transform, macroblock type selection, quantization, rate control and the like. Digital signal processor 108 transfers the compressed data to the bitstream processor 112. Digital signal processor 108 further processes the compressed frame, thus reconstructing the reference frames, and transfers the reconstructed reference frames to the external memory unit via memory controller 110, thereby overwriting some of the existing reference frames.

[61] Bitstream processor 112 encodes the compressed video data into a standard MPEG-1 and MPEG-2 format, in accordance with a sequence known in the art of encoding commands. Bitstream processor 112 transfers compressed video data streams to multiplexing processor 114.

[62] Audio encoder 113 is a processor responsible for audio encoding. In an embodiment of the present invention, audio encoder 113 supports MPEG-1 Layer II and Dolby AC-3 encoding and may be reprogrammed to support various additional audio compression schemes. The audio encoder 113 is also responsible for acquiring the uncompressed audio

signal (I2S and AC97 standards are supported, for example) and buffering the compressed audio. Audio encoder 113 supports encoding of audio signal with a different input sample rate and with a different output bitrate.

[63] Multiplexing processor 114 multiplexes the encoded video and the encoded audio and/or user data streams (as received from bitstream processor 112 and audio encoder 113) and generates, according to a sequence of optimized multiplexing commands, MPEG-2 standard format streams such as packetized elementary stream, program stream, transport stream and the like. Multiplexing processor 114 transfers the multiplexed video/audio/data streams to a compressed data stream output and to memory controller 110. Multiplexing processor 114 outputs a stream of encoded video and/or audio and/or data.

[64] Global controller 104 controls and schedules the video input buffer 102, the motion estimation processors 105 and 106, the digital signal processor 108, the memory controller 110, the bitstream processor 112, the I2C/GPIO interface, and the multiplexing processor 114. Global controller 104 is a central control unit that synchronizes and controls all of the internal chip units and communicates with all of the internal chip units using data-instruction-device buses.

[65] In an embodiment of the present invention, the I2C/GPIO interface 116 may be used to program an external video A/D or an external audio A/D through the single chip device 100. In an embodiment of the

present invention, the I2C/GPIO interface 115 is configured (programmed) through the host interface or global controller 104 using microcode. Fig. 7 illustrates a block diagram of the I2C/GPIO interface 116 in accordance with an embodiment of the present invention.

[66] An embodiment of the present invention provides a digital video broadcasting (DVB) formatter 117 as part of the mux processor 114. The DVB formatter 117 enables an encoded multiplexed stream to be converted to a standard DVB format and transmitted directly from the device 100 to another chip without going through a host interface or PCI interface. The host processor does not need to get involved in the transfer of the encoded data when the DVB interface is used. The DVB interface provides a powerful and smaller interface to transfer encoded data to, for example a decoder chip.

[67] Fig. 8 is a block diagram and timing diagram illustrating the signals and timing output by the DVB formatter 117 in accordance with an embodiment of the present invention. Fig. 8 illustrates a typical system for parallel transmission of a transport stream at either constant or variable rate. The clock (CLOCK), the 8-bit data (Data), and the PSYNC signal are transmitted in parallel. The PSYNC signal marks the sync byte of the transport header and is transmitted each 188 bytes. The DVALID signal is a constant 1 in the 188-byte mode. All signals are synchronous to the clock which is set to the transport bit rate and number of bits. The DVB interface has optional special input signal (STALL), which may be

used by DVB receiver in order to slow and/or stop the transmitter. DVB formatter may work in a master (CLOCK is generated by transmitter) or slave mode (the CLOCK is received).

[68] An embodiment of the present invention provides a vertical blanking interval (VBI) extractor 103 as part of the VIB 102. In general, analog video data may contain user data such as closed caption information or other user information. For example, a CCIR 656 video signal may typically contain uncompressed video data in a picture interval and user data in a VBI interval. The user data is transmitted during the VBI of the video signal where picture data is not present.

[69] The VBI extractor 103 in the VIB 102 extracts the user data from the VBI of the CCIR 656 video stream. The extracted user data is then using microcode in either the mux processor 114 and inserted into the encoded stream or is processed using microcode in the global controller 104 or BSM 112 and inserted in the encoded stream.

[70] Fig. 9 illustrates how the VBI extractor 103 may extract user data from specified lines of a video signal in accordance with an embodiment of the present invention. Several modes may be supported by the VBI extractor 103 and subsequent processing, including a generic VBI mode. In the generic VBI mode, the user defines which pels of which video lines (e.g. of line 6 through line 21) of each field (top, bottom) are to be extracted and further transmitted in the compressed stream and/or for slicing. The VBI extractor 103 may also extract already sliced data (the

data is sliced by external video decoder and inserted to CCIR656 stream) accordingly to SAA7113/SAA7114/SAA7115 format and Ancillary format (SMPTE 291M). The VBI extractor 103 may also extract the data from HBI (horizontal blanking interval). The VBI extractor 103 may be programmed to extract the data also from picture interval for full field slicing.

[71] Several registers are used to control the VBI extractor 103. A first register determines the video lines of the top field to be extracted in generic VBI mode. Each bit of the first register corresponds to a certain video line (see Fig. 9). Through setting the bits of the first register, the user selects the video lines of the top field to be extracted.

[72] A second register determines the video lines of the bottom field to be extracted in generic VBI mode. Each bit of the second register corresponds to a certain video line (see Fig. 9). Through setting the bits of the second register, the user selects the video lines of the bottom field to be extracted.

[73] A third and fourth register determine the pixel interval within a video line of the top field of each frame to be extracted and transmitted in the compressed stream and/or for slicing. The content of the third and fourth registers may range from 0 to 1023, and a START value must be less than an END value.

[74] A fifth and sixth register determine the pixel interval within a video line of the bottom field of each frame to be extracted and transmitted in

23

the compressed stream. The content of the fifth and sixth registers may range from 0 to 1023, and a START value must be less than an END value.

[75] The various elements of device 100 may be combined or separated according to various embodiments of the present invention.

[76] Also, the various elements may be implemented as various combinations of programmable and non-programmable hardware elements.

[77] In summary, certain embodiments of the present invention afford an approach to perform video and audio encoding on a single chip to generate a stream of encoded video and audio data for use in various applications such as personal video recorders, DVD recorders, and set top box recorders. In other words, the system of the present invention enables a single chip that encodes video and audio (and any other system data desired) and generates therefrom a stream of encoded data.

[78] While the invention has been described with reference to certain embodiments, it will be understood by those skilled in the art that various changes may be made and equivalents may be substituted without departing from the scope of the invention. In addition, many modifications may be made to adapt a particular situation or material to the teachings of the invention without departing from its scope. Therefore, it is intended that the invention not be limited to the particular embodiment disclosed, but that the invention will include all embodiments falling within the scope of the appended claims.